Sherry A. Terrell, Ph.D.
NISB/Dres/or/Ord/HCFA
6325 Security Blvd. 2B140M
Baltimore, MD 21207

A RAND NOTE

A Description of Expensive and Long-Staying Patients

Sally Trude, Grace M. Carter

October 1989

REPORTS
RA
971
.3
T782
1989

RAND

This research was sponsored by the Health Care Financing Administration, U.S. Department of Health and Human Services under Cooperative Agreement No. 99-C-98489/9.

The RAND Publication Series: The Report is the principal publication documenting and transmitting RAND's major research findings and final research results. The RAND Note reports other outputs of sponsored research for general distribution. Publications of The RAND Corporation do not necessarily reflect the opinions or policies of the sponsors of RAND research.

Published by The RAND Corporation
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90406-2138

RA
971.3
.T782
1989

Sherry A. Terrell, Ph.D
N1SB/Dres/or/Ord/HCFA
6325 Security Blvd. 2B140M
Baltimore, MD 21207

A RAND NOTE

N-2998-HCFA

A Description of Expensive and Long–Staying Patients

Sally Trude, Grace M. Carter

October 1989

Prepared for
The Health Care Financing Administration,
U.S. Department of Health and Human Services

RAND

RAND/UCLA/HARVARD
Center for Health Care
Financing Policy Research

PREFACE

This note was prepared as part of the RAND/UCLA/HARVARD Health Care Financing Administration's research on outlier payment policy. The research was supported under Cooperative Research Agreement C-98489/9-04. A related publication is Emmett B. Keeler, Grace M. Carter, and Sally Trude, *Insurance Aspects of DRG Outlier Payments*, The RAND Corporation, N-2762-HHS, 1988.

SUMMARY

Some patient groups are more apt to have high costs or long stays relative to the typical value for the diagnosis related group (DRG) to which they are assigned. The identification of expensive and long-staying patient groups highlights the necessity of outlier payments under Medicare's Prospective Payment System. To the extent that patient attributes are associated with costlier care, costlier care appears to relate to patients' needs rather than entirely to the hospital's inefficiency.

If certain types of patients are more apt to be costlier and become outliers, one would want the outlier payment policy to reimburse hospitals adequately for the patients' care. One would also want outlier payment policy to improve payments for these costlier patient groups. Having identified costlier and longer-staying patient groups, one could then evaluate the impact of changes in outlier policy on these patient groups.

In this study, we considered patient groups characterized by patient demographics, chronic disease conditions, number of diseased body systems, and prior hospitalizations. Patients were considered to have extremely long stays if they were among the 5 percent longest stays in their DRGs. Patients were considered to be very expensive if the estimate of the hospital's profit for the case was among the 5 percent largest losses across all DRGs.

We used patient characteristic information in ordinary least squares regression models to describe the 5 percent of cases with the absolute largest losses, the 5 percent of longest stays within each DRG, and a measure of accounting profit for the case. We ran each of these three models separately for cases in surgical DRGs and in medical DRGs. For surgical cases, we found that the likelihood of an extremely large loss increased most for patients who had a secondary condition of pulmonary edema, malnutrition, lung cancer, or chronic liver disease; had multiple system involvement; or had a prior expensive or long stay

in the previous six months. For medical cases, the likelihood of an extremely large loss increased most for patients who had a secondary condition of pulmonary edema, malnutrition, or leukemia; had multiple system involvement; or had a prior expensive or long stay in the previous six months. However, patients in surgical DRGs with these characteristics had a greater likelihood of being extremely expensive cases than did patients in medical DRGs. That is, although malnutrition increases the likelihood of being an expensive case for both surgical and medical cases, a surgical patient with malnutrition is more likely to become an extremely expensive case than is a medical patient.

The likelihood of an extremely long stay (relative to the DRG) increases most for patients with a secondary condition of pulmonary edema and malnutrition, with a prior expensive or long stay, or with multiple diseased body systems. This holds true for both medical and surgical cases. For cases in surgical DRGs, secondary conditions of lung cancer and depression also increase the likelihood of an extremely long stay.

The inclusion or exclusion of sets of explanatory variables in regression models can affect the coefficients for other variables in the model. In this analysis, we found collinearity between the chronic disease and multiple system variables that substantially affected their regression coefficients. The effects of many individual chronic conditions could be captured by the counts of the number of involved body systems. Three chronic conditions—pulmonary edema, malnutrition, and liver disease—increase the chance of a very costly case beyond the increase caused by the additional involved body system.

If some of these patient groups tend to use hospitals that on average incur more losses, the patient characteristic will pick up an effect that is actually attributable to the hospital. To control for such an effect, we ran additional regression models that included a dummy variable for each hospital. After controlling for the hospital, few noticeable changes to the coefficients in the regression equations appeared. After controlling for the hospital, only race had substantially less explanatory power.

ACKNOWLEDGMENTS

The authors would like to thank Katherine L. Kahn, M.D., for identifying the secondary conditions analyzed in this Note. We would also like to thank Daniel Byrne for constructing the analytic data file and C. Richard Neu for helpful comments on an earlier draft. We are especially indebted to Richard L. Kravitz, M.D., for his thoughtful and discerning comments as the reviewer for the final draft. His comments led to further analysis that substantially improved our Note. We would also like to thank Nancy Elwood for preparing this Note.

CONTENTS

PREFACE iii

SUMMARY v

ACKNOWLEDGMENTS vii

TABLES xi

Section

I. INTRODUCTION 1

II. DATA 5
  Measures of Extreme Stays 7
  Patient Demographics 8
  Chronic Disease Indicators 9
  Multisystem Disease Variables 10
  Prior Utilization Variables 14

III. RESULTS 16
  Summary Statistics 16
  Multivariate Model of Patient Characteristics 22
  Comparison of Extremely Long Stays and Extremely Large Losses 26
  Effects of Age and Disability 26
  Collinearity of Multiple Diseased Body Systems and Chronic Diseases 28
  Controlling for Hospital 30

IV. DISCUSSION 38

Appendix: COMPARISON OF 1988 AND 1986 DIAGNOSIS RELATED GROUPINGS 43

REFERENCES 47

TABLES

2.1. Secondary Chronic Disease Specifications 11
2.2. Ranges of Codes Associated with Body System 14
3.1. Means for Analysis Variables by Type of Extreme Indicator, Surgical DRGs 17
3.2. Means for Analysis Variables by Type of Extreme Indicator, Medical DRGs 18
3.3. Mean and Standard Deviation of Profits and Standardized Length of Stay by Patient Characteristics, Surgical DRGs... 19
3.4. Mean and Standard Deviation of Profits and Standardized Length of Stay by Patient Characteristics, Medical DRGs ... 20
3.5. Patient Characteristic Models, Surgical DRGS 23
3.6. Patient Characteristic Models, Medical DRGS 24
3.7. Combined Effect of the Age Coefficients, Surgical DRGs 27
3.8. Comparison of Age Contribution to Age 65, Surgical DRGs ... 29
3.9. Comparison of Age Contribution to Age 65, Medical DRGs 29
3.10. Patient Characteristic Models for 5 Percent Largest Losses, Surgical DRGs 31
3.11. Patient Characteristic Models for 5 Percent Largest Losses, Medical DRGs 32
3.12. Patient Characteristic Models for 5 Percent Largest Losses with and without Control Variables for Each Hospital, Surgical DRGS 34
3.13. Patient Characteristic Models for 5 Percent Largest Losses with and without Control Variables for Each Hospital, Medical DRGS 35
3.14. Patient Characteristic Models for 5 Percent Longest Stays within DRG with or without Control Variables for Each Hospital, Surgical DRGS 36
3.15. Patient Characteristic Models for 5 Percent Longest Stays within DRG with or without Control Variables for Each Hospital, Medical DRGS 37
A.1. Profit 44
A.2. Five Percent Largest Losses 46

I. INTRODUCTION

Under Medicare's Prospective Payment System (PPS), inpatient hospital stays are reimbursed a predetermined price according to the diagnosis related group (DRG) into which the patient's illness falls. Although the cost for any given case may be more or less than the predetermined price, the payments are intended to cover the hospital's costs on average.

To limit the risk to hospital finances associated with extremely long or expensive stays, Medicare supplements the predetermined payment with an outlier payment. Cases that exceed thresholds of length of stay or cost are paid additional amounts based on how far these atypical cases exceed the thresholds. These payments are intended to reduce the hospital's losses from outlier cases, not to eliminate the losses. The remaining losses from these cases, like losses on nonoutlier cases, are intended to be offset by profits on other cases.

Although hospitals receive outlier payments, outlier cases still cause losses to hospitals. Hospitals, therefore, have incentives to discourage admissions of potential outliers and other cases that cause losses. If certain patient groups are identifiable as unprofitable before admission, hospitals may persuade physicians to admit them elsewhere or may choose not to provide special resources unprofitable patient groups require.

This study identifies expensive and long-staying patient groups. We consider patient demographics, chronic disease conditions, number of diseased body systems, and prior hospitalizations. Although some of these characteristics may not be apparent to a physician evaluating a patient for the first time in an emergency room, they are apt to be familiar to the patient's primary-care physician or to the physician to whom the patient has been referred. This study does not determine whether hospitals actually do avoid unprofitable patients or expensive patient groups. We can only conjecture as to possible mechanisms by which hospitals might deny access, and suggest ways in which to study whether these patient groups do face limited access to care.

The identification of costlier patients underscores the importance of an outlier policy. The outlier payment policy can be viewed either as rewarding inefficient hospitals or as paying hospitals that care for sicklier–and hence costlier–patients. To the extent that patient attributes are associated with costlier care, costlier care appears to relate to the needs of patients rather than entirely to the hospital's inefficiency.

In this paper, we simulate a policy of no outlier payment. We then identify those patient groups that represent the largest losses or the longest stays. If certain types of patients are more apt to be costlier and become outliers, one would want the outlier payment policy to reimburse hospitals adequately for care of these patients. One would also want outlier payment policy to improve payments for these costlier patient groups. Having identified costlier and longer-staying patient groups, one could then evaluate the impact of changes in outlier payment policy on these patient groups. For an example of how changes in outlier payment policy impact payments for different patient groups, see Keeler et al. (1988).

Current outlier policy distinguishes between length of stay (day) outliers and cost outliers, paying them differently. We will describe the outlier policy in effect during fiscal year 1988. The policy was similar in earlier years, although some parameters were slightly different. Day outliers are defined as those hospital stays that exceed the mean length of stay for the same DRG by the lesser of 18 days or two standard deviations. The outlier payment for these cases is 60 percent of the per diem federal payment rate for the DRG for the days beyond the outlier threshold. If a stay is not a day outlier, but the charges for the care exceeds the greater of two times the federal payment rate for the DRG or $14,000, the stay is classified as a cost outlier. The cost outlier payment is 60 percent of the additional cost beyond the cost outlier threshold. A stay is paid as a day outlier if it exceeds the day outlier thresholds, even if it also exceeds the cost outlier threshold.

Significant changes to the outlier policy became effective as of November 1, 1988 (*Federal Register,* September 30, 1988). Cost outlier payments were increased to 75 percent of the additional cost of the stay beyond the cost threshold.[1] In addition, the length of stay threshold no longer takes precedence. Instead, hospitals are paid the larger of the day outlier payment or the cost outlier payment. This policy helps to compensate longer stays that also have extremely high costs. The thresholds also change, so that the threshold for day outliers increases to the mean length of stay plus 24 days or three standard deviations, and the cost outlier threshold increases to the greater of $28,000 or two times the federal payment rate.

In a study investigating outlier policy, Keeler et al. (1988) found a substantial number of cases that exceeded the cost outlier threshold but still incurred enormous losses for hospitals because the hospitals were paid only a small day outlier payment. If outlier payment policy fails to reimburse adequately for the care of costly patient groups, these groups may face limited access to care, or the hospitals that care for these groups may suffer large losses.

We investigate patient attributes that predict longer stays and that predict higher costs. Early PPS legislation incorporated length of stay as a proxy for cost and gave precedence to length of stay outliers. Recent research by the Congressional Research Service (1987) and Keeler et al. (1988) has shown that length of stay is not a good proxy for cost. Because regional differences in length of stay exist, some people have argued in favor of continued reliance on length of stay outliers to avoid severe changes in the distribution of Medicare monies. In addition, it may be true that long-staying patients are visible and may be inappropriately discharged. Therefore, this paper addresses both expensive and long-staying patients.

This Note first describes the data sources for the analysis and describes the patient characteristics and measures of atypical cases in detail. The next section describes the multivariate models of patient characteristics, compares cases with long stays to those with extreme losses, and discusses the relationship between the hospital and patient characteristics. The final section discusses how the results of the analysis relate to our policy concerns.

---

[1] Except for burn cases, which receive 90 percent of cost beyond the threshold.

II. DATA

Medicare claims for inpatient acute care hospital stays were the main source of data for this analysis. An inpatient claim is processed for each Medicare beneficiary's hospitalization. For research purposes, the Health Care Financing Administration (HCFA) selects a 20 percent random sample of beneficiaries and creates a data file containing any hospital claims for these beneficiaries. Discharges occurring between July 1, 1984, and June 30, 1985, for this 20 percent random sample of beneficiaries were used as our primary analysis sample. Discharges occurring after October 1, 1983, but before July 1, 1984, were used with the primary analysis sample to measure prior hospital use. We excluded from the analysis claims from the four waiver states (Maryland, Massachusetts, New Jersey, and New York).

The hospital stay was the unit of analysis. If a beneficiary had more than one hospital stay during the analysis year, each hospital stay was included. The data do not represent beneficiaries who were not hospitalized during the year.

These Medicare claims provide basic information about the hospital stay's characteristics, such as the principal diagnosis, up to four secondary diagnoses, surgical procedures, dates of admission and discharge, and total number of days and charges covered by Medicare payments. The claims also contain information describing the patient, such as age, race, sex, and end-stage renal disease status.

Each hospital claim has sufficient information for categorizing the hospital stay into a DRG. For each claim, both the 1986 and 1988 versions of the Grouper program were used to assign DRGs. Significant changes were made to the algorithm between 1986 and 1988, especially in terms of how cases were assigned to DRG "pairs." Some DRGs are paired so that both members of the pair contain cases with the same principal diagnosis and procedures but one of the pair contains cases with complications and comorbidities. For example, DRG 300 contains endocrine disorders with complications and comorbidities and DRG 301 contains endocrine disorders without complications and comorbidities. The DRG that includes complications and comorbidities has a higher weight, and hence receives a higher average price. The existence of complications and comorbidities is determined by the coding of specific additional secondary conditions. In the 1986 version of Grouper, all allowable secondary diagnostic codes would upcode to a DRG with complications and comorbidities. Any diagnosis on the list would result in a higher classification for all the DRG "pairs." With the 1988 version of Grouper, some secondary diagnoses no longer apply to some principal diagnoses. For example, a secondary diagnosis of adult-onset diabetes with renal manifestations (25040) is no longer considered a complication or comorbidity for a primary diagnosis of hypoglycemic coma (2513x). By restricting the conditions for which a disease qualifies as a complication or comorbidity, redundancy in coding is reduced. A complication or comorbidity is now more likely to be a separate disease entity rather than an extension of the principal diagnosis. The 1988 Grouper makes an additional change by eliminating age 70 and over as a complication.

We initially conducted our analysis using the 1986 Grouper. When the 1988 Grouper became available, the changes seemed sufficiently important to this analysis to rerun it using the new DRGs. Indeed, a substantial difference in the results using the two versions occurred. The appendix details the comparison of the results using the two Grouper versions. In the main body of this Note, we restrict our attention to the results based on the 1988 Grouper.

We developed five basic groups of variables for the analysis: measures of extreme stays, patient demographics, chronic disease indicators, multisystem disease variables, and prior utilization variables.

An extreme stay is a hospital stay that is either exceptionally long or exceptionally expensive. Patient demographics include age, sex, and race. Chronic disease indicators identify patients with any of 19 chronic disease conditions, such as chronic obstructive lung disease, hypertension, or chronic liver disease. The multisystem disease variables count the number of different diseased body systems identified by the diagnostic codes. The prior utilization variables identify patients with hospitalizations that occurred within six months of the current hospital stay. The five subsections that follow describe in further detail the derivation of each type of variable.

MEASURES OF EXTREME STAYS

We developed two measures of extreme stays. One indicates hospital stays that are exceptionally long (LOUT05); the other indicates stays that cause large losses for the hospital (PROUT05).

To identify an extremely long hospital stay, the total number of days covered by Medicare payments for the hospital stay was taken from each claim. Since the extremity of a length of stay is relative to the DRG, we standardized by DRG the total covered days for each claim. We subtracted the log of the geometric mean length of stay for the DRG from the log of the total number of covered days, then divided the difference by the standard deviation of the log of the total covered days.[1] We obtained the geometric mean length of stay for each DRG from the *Federal Register* (Sept. 1, 1987). We derived the standard deviation for each DRG from the hospital claims analysis file.

Once we developed the standardized lengths of stay (STDLOS), we could inspect the distribution. The mean of the standardized distribution was -.005, with a standard deviation of 0.965.[2] From this distribution, we identified extremely long stays in the tail (LOUT05). For this analysis, we used a cutoff of the 5 percent longest stays, which corresponds to a standardized value exceeding 1.5.

To identify hospital stays that represent large losses for the hospital, we developed a measure of accounting profit. First, we adjusted the total covered charges for the case by multiplying by the hospital's 1984 cost-to-charge ratio. We calculated the cost-to-charge ratios using PPS I cost report data and departmental charges from a 20 percent sample of PPS I Medicare bills. (PPS I refers to the first 12-month period each hospital was on PPS.) Second, we simulated payments as if PPS were fully implemented but no outlier payments existed. We did not simulate hospital-specific payment rates. The simulated payment rates were budget neutral to the fiscal year 1985 federal rates. To produce budget neutrality while excluding outlier payments, we increased the fiscal year 1985 federal payment rates to produce exactly the same total dollar payments.[3] For discharges occurring before fiscal year 1985, we decreased the federal rates by the market basket inflation rate to adjust for inflation. For all cases, we adjusted the federal rate to include the indirect medical education payments for teaching hospitals. Rural referral centers received urban federal payments. We made no adjustment for disproportionate-share hospitals. Finally, we subtracted the total covered charges—adjusted to an estimate of cost for the case—from the estimated hospital payment to obtain a measure of profit for the hospital stay.

---

[1] The geometric mean for a DRG's length of stay is obtained by taking the average of the log of total covered days for a DRG and exponentiating the average to return the value to the original unit in days. Length of stay has approximately a log normal distribution, so this reduces the influence of extremely long lengths of stay on the mean.

[2] Because we used the geometric mean length of stay from the *Federal Register* instead of computing it from our analysis sample, the mean and standard deviation are not identical to zero and one, respectively.

The distribution for this profit variable (PROFIT) has a mean of $532 and a standard deviation of $3443. From the distribution of profits, a 5 percent cutoff for extremely large losses was determined to be at losses of $4000 or more (PROUT05). In this analysis, we used both the indicator for large losses and the profit variable.

PATIENT DEMOGRAPHICS

The only patient demographics available in the hospital claims are age, sex, and race. For the analysis, we used the natural log of age (LAGE). To capture differences between the aged and the disabled populations, we included an indicator for beneficiaries under the age of 65. To capture the separate effects of aging on these two population groups, we included an interaction term (log of age multiplied by the disability indicator). An indicator for sex identifies male beneficiaries (MALE), and an indicator for race distinguishes black beneficiaries (BLACK).

---

[3] We multiplied urban and rural federal payment rates by the same proportion. Outlier payments constitute approximately 4 percent of federal payments in our analysis sample.

CHRONIC DISEASE INDICATORS

A wide variety of chronic disease conditions can make a hospital stay costlier or longer. Some of these conditions are specifically considered in determining the payment for the hospital stay. That is, in the case of DRG "pairs," certain chronic conditions are considered complications and comorbidities and result in a higher payment. Some chronic conditions this analysis examines are considered complications and comorbidities, while others are not.

An internist provided us with a list of secondary conditions that, in her clinical judgment, would lead to an extended hospital stay. Her list provided both acute and chronic conditions. We selected from her list those conditions that had a chronic component.

Our next step involved noting each diagnostic code listed under the relevant disease heading in the alphabetic index (Volume 2) of the *International Classification of Disease, 9th Revision, Clinical Modification* (ICD-9-CM) manual. For example, under the term *anemia* are several pages of disease conditions and their associated diagnostic codes. We excluded conditions that were obviously acute or not pertinent to the Medicare population (for example, fetal, or complicating pregnancy). We then found the resulting list of diagnostic codes in the tabular list of diseases. From this list we determined the level of coding sufficient to capture the disease condition. For example, the three-digit level code 493 for asthma was sufficient to describe that disease. However, depression is identified by four four-digit level codes and one three-digit code. Table 2.1 shows the list of diseases and their associated diagnostic codes that define the chronic disease indicators. It also identifies the associated variable name used in subsequent tables.

Each hospital claim includes a principal diagnosis and up to four secondary diagnoses. If any of the four secondary diagnoses represented a chronic condition unrelated to the principal diagnosis, we identified the record with the chronic condition. If the secondary diagnosis coded on the claim assigned the case to the same DRG it would have had it been the principal diagnosis, we did not consider that secondary diagnosis to be a separate chronic condition.

MULTISYSTEM DISEASE VARIABLES

The multisystem disease variables count the number of diseased body systems the hospital claim describes. These variables are designed as a simple measure of case complexity using the data readily available within the hospital claim. This severity measure is similar to the measure developed by Mendenhall (1984), which is based on assigning diagnoses to major diagnostic categories (MDCs) and then counting the number of distinct MDCs. Using the ICD-9-CM classification, we identified nine distinct body systems. Table 2.2 lists these systems with the associated ranges for the ICD-9-CM codes. We did not include some major categories in the ICD-9-CM classification system for counting diseased body systems. Some categories—such as infections and parasites, or injury and poisoning—are more pervasive and not confined to a single body system. Neoplasms are also categorized separately by the ICD-9-CM system. For example, leukemia is classified under neoplasms and is not considered a disease of the blood. As a result, neither leukemia nor lung cancer would be included in a count of diseased body systems. We excluded other categories, such as complications of pregnancy, if they did not pertain to the Medicare population. Because of these exclusions, a patient could potentially be identified as having no body system involvement. The maximum possible number of involved body systems is nine.

We use both the principal diagnosis and all coded secondary diagnoses in each hospital claim for counting the number of diseased body systems. A patient with three respiratory conditions and one musculoskeletal disorder would have two diseased body systems.

For the purposes of the analysis, we wanted to identify the additional increment in cost or length of stay associated with each additional involved body system. By developing indicator variables that identify each additional involved body system, we could then use

Table 2.1

SECONDARY CHRONIC DISEASE SPECIFICATIONS

Variable Name	ICD-9-CM Codes	Disease
OBSTRC	492xx	Emphysema
	496xx	Chronic airway obstruction, not elsewhere classified
ASTHMA	493xx	Asthma
LFIB	515xx	Postinflammatory pulmonary fibrosis
DEPRES	2963x	Major depressive disorder, recurrent episode
	2980x	Depressive-type psychosis
	3004x	Neurotic depression
	3091x	Prolonged depressive reaction
	311xx	Depressive disorder, not elsewhere classified
LNGCAN	1622x	Malignant neoplasm of main bronchus
	1623x	Malignant neoplasm of upper lobe, bronchus or lung
	1624x	Malignant neoplasm of middle lobe, bronchus or lung
	1625x	Malignant neoplasm of lower lobe, bronchus or lung
	1628x	Malignant neoplasm of other parts of bronchus or lung
	1629x	Malignant neoplasm of bronchus and lung, unspecified
	1970x	Malignant neoplasm of lung
ATHERO	4471x	Stricture of artery
	4350x	Basilar artery syndrome
	4351x	Vertebral artery syndrome
	4358x	Other specified transient cerebral ischemias
	4359x	Unspecified transient cerebral ischemia
	436xx	Acute but ill-defined cerebrovascular disease
	437xx	Other and ill-defined cerebrovascular disease
	438xx	Late effects of cerebrovascular disease
	4439x	Peripheral vascular disease, unspecified
	5571x	Chronic vascular insufficiency of intestine

Table 2.1 (Continued)

Secondary Chronic Disease Specifications

Variable Name	ICD-9-CM Codes	Disease
PULEDM	514xx	Pulmonary congestion and hypostasis (chronic)
LEUK	2024x	Leukemic reticuloendotheliosis
	204xx	Lymphoid leukemia
	205xx	Myeloid leukemia
	206xx	Monocytic leukemia
	207xx	Other specified leukemia
	208xx	Leukemia of unspecified cell type
MALNUT	260xx	Kwashiorkor
	261xx	Nutritional marasmus
	262xx	Other severe protein-calorie malnutrition
	263xx	Other and unspecified protein-calorie malnutrition
	7994x	Cachexia
OBESE	2780x	Obesity
ANEMIA	280xx	Iron deficiency anemias
	281xx	Other deficiency anemias
	283xx	Acquired hemolytic anemias
	284xx	Aplastic anemia
	285xx	Other and unspecified anemias
HHANEMIA	282xx	Hereditary hemolytic anemias
ARTHR	2740x	Gouty arthropathy
	6960x	Psoriatic arthropathy
	711xx	Arthropathy associated with infections
	713xx	Arthropathy associated with other disorders classified elsewhere
	714xx	Rheumatoid arthritis and other inflammatory polyarthropathies
	715xx	Osteoarthrosis and allied disorders
	716xx	Other and unspecified arthropathies
	7193x	Palindromic rheumatism
	7210x	Cervical spondylosis without myelopathy
	7212x	Thoracic spondylosis without myelopathy
	7213x	Lumbosacral spondylosis without myelopathy
	7218x	Other allied disorders of spine
	7219x	Spondylosis of unspecified site

Table 2.1 (Continued)

Secondary Chronic Disease Specifications

Variable Name	ICD-9-CM Codes	Disease
DIAB2	25000	Diabetes mellitus, adult onset, no complications
DIAB3	25001	Diabetes mellitus, juvenile onset, no complications
DIAB4	250x0	Diabetes mellitus, adult onset, complications
DIAB5	250x1	Diabetes mellitus, juvenile onset, complications
HYPER	401xx	Essential hypertension
DEMENT	0941x	General paresis
	290xx	Senile and presenile organic psychotic conditions
	2912x	Other alcoholic dementia
	2941x	Dementia in conditions classified elsewhere
PSYCH	2950x	Simple type schizophrenic disorder
	2951x	Disorganized type schizophrenic disorder
	2952x	Catatonic type schizophrenic disorder
	2953x	Paranoid type schizophrenic disorder
	2989x	Unspecified psychosis
LIVER	571xx	Chronic liver disease and cirrhosis
	572xx	Liver abscess and sequelae of chronic liver disease
RENDUM	(a)	Medicare end-stage renal disease patient, not in renal major diagnostic category (302 <= DRG <= 333)

(a) Not applicable.

Table 2.2

RANGES OF CODES ASSOCIATED WITH BODY SYSTEM

Codes	Body System
240-279	Endocrine, nutritional, and metabolic diseases and immunity disorders
280-289	Diseases of the blood and blood-forming organs
320-389	Diseases of the nervous system and sense organs
390-459	Diseases of the circulatory system
460-519	Diseases of the respiratory system
520-579	Diseases of the digestive system
580-629	Diseases of the genitourinary system
680-709	Diseases of the skin and subcutaneous tissue
710-739	Diseases of the musculoskeletal system and connective tissue

regression coefficients for each indicator variable to represent the additional increment associated with the additional involved body system. The variable DMULT1 identifies patients with involvement of more than one body system and DMULT2 identifies patients with involvement of more than two body systems. Therefore, anyone with DMULT2=1 will also have DMULT1=1.

PRIOR UTILIZATION VARIABLES

Patients who had prior hospitalizations seemed potentially to be sicker, and therefore at greater risk of more expensive or longer stays. Prior utilization variables are intended to capture this increased risk.

Our primary analysis sample consists of discharges in the year from July 1, 1984, to June 30, 1985. However, claims data before this time period were also available. We decided to look at hospitalizations that occurred within six months before the current stay.

For each discharge occurring between July 1, 1984, and June 30, 1985, we noted the date of admission. We counted as a prior stay any admission that had occurred six months before the current admission. The total number of prior stays—and whether or not any of the prior stays had been exceptionally expensive or long-was appended to the current stay. We considered a prior stay exceptionally expensive or long if the loss or standardized length of stay was in the 5 percent tail of its distribution (POUT05). (The algorithm for identifying prior stays that are exceptionally expensive or long mirrors the computation of extreme stays discussed earlier.)

For the regression models, we converted the count of prior hospitalizations to a series of additive indicators. NSTAY1 indicates cases with one or more prior hospitalizations, NSTAY2 indicates cases with two or more prior hospitalizations, and so on. Any case indicated by NSTAY2 will also be indicated by NSTAY1. As we would expect, as the number of prior hospitalizations increases, the number of cases indicated declines.

III. RESULTS

This section reviews the results of our analysis, beginning with summary information for each variable. It then describes three ordinary least squares (OLS) regressions which we ran separately for cases in surgical and medical DRGs: the 5 percent longest stays within each DRG, the 5 percent largest losses, and the accounting profit. We discuss individually the results of each model. We then compare the patient attributes that best predict long stays to the patient attributes that best predict expensive stays. Because the Medicare population includes both disabled and aged populations, we provide a separate discussion of the effects of disability status, age, and their interaction. Next, we describe the collinearity of the multiple diseased body system variables with the chronic disease indicators. This section ends with a discussion of the results of the OLS regressions after controlling for the effect of individual hospitals. Each set of analyses considers cases in surgical DRGs separately from those in medical DRGs.

SUMMARY STATISTICS

Tables 3.1 and 3.2 show the means for each variable in the analysis. Table 3.1 provides means for cases in surgical DRGs; Table 3.2 provides means for cases in medical DRGs. The first column shows the average for all cases in the primary analysis file. The second column shows the mean of each variable for cases in the 5 percent tail of the largest losses. The third column shows the mean of each variable for cases in the 5 percent tail of the longest stays relative to their DRG. As described earlier, the average profit across all cases is $532; the average standardized length of stay is -.005. For surgical cases (Table 3.3), the average profit is $735; the average standardized length of stay is 0.075. For medical cases (Table 3.4), the average profit is $458; the average standardized length of stay is -.034.

Table 3.1

MEANS FOR ANALYSIS VARIABLES BY TYPE OF EXTREME INDICATOR, SURGICAL DRGS

Variable	All Surgical Cases	Extreme Loss	Extreme Length of Stay
PROFIT	735.1317	-11602.8713	-7343.6891
LOUT05	0.0597	0.4398	1.0000
PROUT05	0.0566	1.0000	0.4172
OBSTRC	0.0603	0.0772	0.0535
ASTHMA	0.0058	0.0051	0.0061
LFIB	0.0027	0.0028	0.0024
DEPRES	0.0042	0.0051	0.0079
LNGCAN	0.0061	0.0111	0.0108
ATHERO	0.0358	0.0510	0.0481
PULEDM	0.0011	0.0054	0.0027
LEUK	0.0020	0.0029	0.0024
MALNUT	0.0044	0.0188	0.0107
OBESE	0.0099	0.0056	0.0073
ANEMIA	0.0412	0.0595	0.0528
HHANEMIA	0.0005	0.0006	0.0004
ARTHR	0.0435	0.0276	0.0470
DIAB2	0.0437	0.0332	0.0452
DIAB3	0.0242	0.0286	0.0247
DIAB4	0.0102	0.0142	0.0138
DIAB5	0.0135	0.0259	0.0185
HYPER	0.1234	0.0742	0.0973
DEMENT	0.0058	0.0050	0.0075
PSYCH	0.0029	0.0037	0.0048
LIVER	0.0067	0.0160	0.0093
RENDUM	0.0145	0.0257	0.0127
UNDER 65	0.0928	0.1143	0.0925
LAGE	4.2756	4.2737	4.2938
UNDER 65*LAGE	0.3639	0.4449	0.3634
MALE	0.4804	0.5183	0.4915
BLACK	0.0703	0.1023	0.1169
DMULT1	0.4615	0.6484	0.5946
DMULT2	0.1659	0.2742	0.2370
DMULT3	0.0289	0.0551	0.0450
DMULT4	0.0017	0.0042	0.0031
POUT05	0.0283	0.0677	0.0522
NSTAY1	0.3100	0.4296	0.3441
NSTAY2	0.0926	0.1599	0.1097
NSTAY3	0.0288	0.0557	0.0330
N	441183	24972	26323

Table 3.2

MEANS FOR ANALYSIS VARIABLES BY TYPE OF EXTREME INDICATOR, MEDICAL DRGS

Variable	All Medical Cases	Extreme Loss	Extreme Length of Stay
PROFIT	457.9966	-8718.2566	-5397.8241
LOUT05	0.0421	0.5239	1.0000
PROUT05	0.0322	1.0000	0.4010
OBSTRC	0.1002	0.1553	0.1046
ASTHMA	0.0101	0.0126	0.0110
LFIB	0.0055	0.0064	0.0054
DEPRES	0.0138	0.0089	0.0195
LNGCAN	0.0172	0.0212	0.0191
ATHERO	0.0492	0.0479	0.0640
PULEDM	0.0033	0.0092	0.0056
LEUK	0.0041	0.0078	0.0060
MALNUT	0.0106	0.0281	0.0235
OBESE	0.0117	0.0057	0.0083
ANEMIA	0.0589	0.0749	0.0780
HHANEMIA	0.0009	0.0010	0.0008
ARTHR	0.0568	0.0239	0.0508
DIAB2	0.0542	0.0349	0.0462
DIAB3	0.0290	0.0307	0.0288
DIAB4	0.0139	0.0144	0.0146
DIAB5	0.0166	0.0239	0.0207
HYPER	0.1197	0.0668	0.0879
DEMENT	0.0120	0.0088	0.0145
PSYCH	0.0073	0.0068	0.0094
LIVER	0.0073	0.0138	0.0099
RENDUM	0.0118	0.0203	0.0104
UNDER 65	0.1122	0.1108	0.1061
LAGE	4.2862	4.2872	4.2996
UNDER 65*LAGE	0.4414	0.4390	0.4202
MALE	0.4407	0.4911	0.4009
BLACK	0.0873	0.1066	0.1175
DMULT1	0.6182	0.7775	0.7273
DMULT2	0.2548	0.3859	0.3508
DMULT3	0.0519	0.0900	0.0820
DMULT4	0.0036	0.0073	0.0070
POUT05	0.0479	0.1141	0.1024
NSTAY1	0.4080	0.4890	0.4741
NSTAY2	0.1706	0.2136	0.1976
NSTAY3	0.0735	0.0889	0.0786
N	1222530	39374	51438

Table 3.3

MEAN AND STANDARD DEVIATION OF PROFITS AND STANDARDIZED LENGTH OF STAY BY PATIENT CHARACTERISTICS, SURGICAL DRGS

		Profit		Standardized Length of Stay	
Variable	N	Mean	SD	Mean	SD
OBSTRC	26,582	850	6419	.025	.971
ASTHMA	2,566	734	5543	.115	.984
LFIB	1,202	977	5710	.054	.943
DEPRES	1,834	567	4683	.404	1.004
LNGCAN	2,672	841	6274	.256	1.078
ATHERO	15,804	777	5364	.206	1.000
PULEDM	465	-2429	9966	.497	1.172
LEUK	881	677	8903	.140	1.029
MALNUT	1,957	-1682	9468	.539	1.019
OBESE	4,383	1117	3369	.059	.908
ANEMIA	18,182	899	5211	.239	.953
HHANEMIA	209	1368	3915	.141	.990
ARTHR	19,206	886	3182	.149	.955
DIAB2	19,301	1081	4018	.130	.959
DIAB3	10,659	1148	4881	.087	.972
DIAB4	4,486	1017	5505	.170	.997
DIAB5	5,946	434	8274	.233	1.002
HYPER	54,460	1188	3914	.055	.924
DEMENT	2,577	1275	3421	.096	1.052
PSYCH	1,285	688	5801	.262	1.025
LIVER	2,948	370	8239	.110	1.099
RENDUM	6,393	1194	6780	-.170	1.074
POUT05	12,483	-209	8476	.335	1.081
NSTAY1	136,767	815	6104	.099	1.002
NSTAY2	40,832	730	6761	.121	1.028
NSTAY3	12,720	648	7171	.122	1.044
UNDER 65	40,920	897	6014	.022	1.010
MALE	211,955	687	5484	.044	1.002
BLACK	31,007	377	6485	.298	1.028
DMULT1	203,606	427	6051	.182	1.003
DMULT2	73,190	170	6632	.234	1.011
DMULT3	12,756	-51	6842	.286	1.014
DMULT4	747	-448	7541	.309	1.079
All Cases	441,183	735	5061	.075	.986

NOTE: SD = standard deviation.

Table 3.4

MEAN AND STANDARD DEVIATION OF PROFITS AND STANDARDIZED LENGTH OF STAY BY PATIENT CHARACTERISTICS, MEDICAL DRGS

		Profit		Standardized Length of Stay	
Variable	N	Mean	SD	Mean	SD
OBSTRC	122,506	185	3078	.015	.944
ASTHMA	12,344	175	2800	.083	.908
LFIB	6,694	451	2802	.045	.917
DEPRES	16,907	420	1895	.190	.911
LNGCAN	21,029	462	2776	-.044	1.016
ATHERO	60,124	361	2457	.080	.966
PULEDM	3,976	-481	3761	.146	1.040
LEUK	5,018	210	3537	-.027	1.090
MALNUT	12,899	-229	4133	.263	1.029
OBESE	14,344	567	1881	-.033	.890
ANEMIA	71,948	339	2754	.127	.941
HHANEMIA	1,057	377	3759	.094	.920
ARTHR	69,474	603	1657	.055	.879
DIAB2	66,235	614	1977	-.003	.912
DIAB3	35,490	528	2615	-.011	.946
DIAB4	16,962	453	2490	.054	.926
DIAB5	20,349	253	2771	.108	.944
HYPER	146,345	660	1964	-.056	.903
DEMENT	14,689	598	1955	.097	.941
PSYCH	8,914	499	2235	.095	.965
LIVER	8,927	111	3374	.089	.987
RENDUM	14,463	495	3789	-.300	1.065
POUT05	58,607	-198	3678	.221	1.071
NSTAY1	498,787	398	2743	-.001	.995
NSTAY2	208,510	388	2793	-.024	1.011
NSTAY3	89,869	424	2719	-.080	1.018
UNDER 65	137,109	489	2661	-.103	.993
MALE	538,775	456	2808	-.102	.984
BLACK	106,767	386	2927	.097	.970
DMULT1	755,707	319	2875	.058	.952
DMULT2	311,449	179	3152	.134	.946
DMULT3	63,419	67	3304	.197	.944
DMULT4	4,434	-111	3872	.269	.951
All Cases	1,222,530	458	2597	-.034	.969

NOTE: SD = standard deviation.

The averages for the indicator variables—chronic disease, multiple diseased body systems and prior utilization—show the prevalence of these conditions. If a condition occurs very frequently in the general Medicare population, a hospital would not be able to use the condition as a predictor of potential outliers. The hospital would target too many patients that would not become outliers. Some 62 percent of the medical cases had two or more diseased body systems; 25 percent had three or more. Approximately 41 percent of the medical cases had had a prior hospitalization in the preceding six months. These conditions are somewhat less common for surgical cases, but still represent a substantial fraction of the cases. Some 46 percent of the surgical cases had two or more diseased body systems; 17 percent had three or more. Approximately 31 percent of the surgical cases had had a prior hospitalization in the preceding six months. Hypertension, occurring in 12 percent of both groups, was the most common secondary disease condition.

Although having three or more diseased body systems was relatively common, only some 5 percent of discharges for medical cases had four or more diseased body systems; 0.4 percent had five or more. Having four or more diseased body systems is even less common than for surgical cases. And although prior hospitalizations were so common, a prior expensive or long stay only occurred in 5 percent of medical cases and in 3 percent of surgical cases. Most secondary chronic disease conditions occur for less than 5 percent of the discharges. Note too that except for end-stage renal disease and hypertension, surgical cases are less apt to have chronic secondary conditions than are medical cases.

Some conditions are uncommon in both the general Medicare population and in categories that represent the largest losses or longest stays. However, their likelihood of occurring in the largest loss or longest-stay groups is much greater. Patients who represent large losses for hospitals are almost three times as likely as the average patient to have pulmonary edema, and twice as likely to have liver disease, a previous expensive hospitalization, or more than four diseased body systems. On the other hand, patients with large losses are less likely to have secondary complications of arthritis, hypertension, or uncomplicated adult-onset diabetes.

Tables 3.3 and 3.4 show the average profit and standardized length of stay for each patient group in surgical DRGs and medical DRGs, respectively. The largest losses on average are shown for the secondary conditions of pulmonary edema and malnutrition, multiple diseased body systems, and a prior long or expensive hospital stay. Patient groups with a secondary condition of liver disease also show a much lower average profit per case than the average for all cases.

Patients with secondary conditions of depression, malnutrition, or pulmonary edema, with prior expensive or long hospitalizations, or with multiple diseased body systems have much longer stays relative to the DRG. Except for the secondary condition of depression, these attributes are also associated with a much lower average profit.

MULTIVARIATE MODEL OF PATIENT CHARACTERISTICS

We used the complete set of patient characteristics—including the demographic, chronic disease, multisystem, and prior utilization variables—in three OLS regressions, each of which we ran separately for medical and surgical DRGs.[1] The first model describes the 5 percent longest stays within each DRG. The second model describes the cases with 5 percent of the absolute largest losses. The third model describes accounting profit for the case. This latter model explains the average loss or gain expected for the patient attributes. Table 3.5 presents the results of the three regression models for surgical cases. Table 3.6 presents the models for cases in medical DRGs.

For surgical cases, the likelihood of an extremely long stay increases by roughly 0.05 (that is, the likelihood roughly doubles) for patients with secondary conditions of depression, lung cancer, pulmonary edema, and malnutrition; it also increases substantially for patients

---

[1] This procedure is equivalent to a Fisher discriminant analysis. Given the size of our sample, the maximum likelihood logit model would have been too costly. Haggstrom (1983) proved that for normally distributed independent variables, OLS regression is equivalent to the maximum likelihood logit model.

Table 3.5

PATIENT CHARACTERISTIC MODELS, SURGICAL DRGS

	5% Longest Stays		5% Largest Losses		Accounting Profit	
Variable	Parameter Estimate	Significance Level	Parameter Estimate	Significance Level	Parameter Estimate	Significance Level
INTERCEP	-0.4523	0.0001	-0.1950	0.0001	7842.316	0.0001
UNDER 65	0.4572	0.0001	0.3071	0.0001	-6918.634	0.0001
LAGE	0.1146	0.0001	0.0522	0.0001	-1608.379	0.0001
UNDER 65*LAGE	-0.1068	0.0001	-0.0724	0.0001	1660.685	0.0001
MALE	0.0035	0.0001	0.0053	0.0001	-88.394	0.0001
BLACK	0.0393	0.0001	0.0228	0.0001	-406.398	0.0001
RENDUM	-0.0281	0.0001	0.0034	0.2917	690.976	0.0001
OBSTRC	-0.0318	0.0001	-0.0187	0.0001	772.581	0.0001
ASTHMA	-0.0116	0.0167	-0.0319	0.0001	433.422	0.0001
LFIB	-0.0268	0.0001	-0.0300	0.0001	782.658	0.0001
DEPRES	0.0528	0.0001	0.0083	0.1399	-185.883	0.1245
LNGCAN	0.0499	0.0001	0.0497	0.0001	-7.833	0.9376
ATHERO	0.0094	0.0001	0.0098	0.0001	255.482	0.0001
PULEDM	0.0641	0.0001	0.1990	0.0001	-2489.230	0.0001
LEUK	0.0060	0.4602	0.0206	0.0109	29.414	0.8660
MALNUT	0.0458	0.0001	0.1382	0.0001	-1619.342	0.0001
OBESE	-0.0262	0.0001	-0.0399	0.0001	651.491	0.0001
ANEMIA	-0.0165	0.0001	-0.0166	0.0001	965.608	0.0001
HHANEMIA	-0.0532	0.0014	-0.0367	0.0259	1324.462	0.0002
ARTHR	-0.0124	0.0001	-0.0436	0.0001	633.765	0.0001
DIAB2	-0.0195	0.0001	-0.0440	0.0001	916.256	0.0001
DIAB3	-0.0222	0.0001	-0.0251	0.0001	1004.894	0.0001
DIAB4	-0.0075	0.0409	-0.0205	0.0001	1002.960	0.0001
DIAB5	-0.0073	0.0248	0.0038	0.2407	418.070	0.0001
HYPER	-0.0296	0.0001	-0.0437	0.0001	860.106	0.0001
DEMENT	0.0021	0.6570	-0.0172	0.0003	761.121	0.0001
PSYCH	0.0307	0.0001	0.0083	0.2126	65.182	0.6499
LIVER	0.0119	0.0084	0.0586	0.0001	-33.370	0.7287
DMULT1	0.0355	0.0001	0.0472	0.0001	-797.808	0.0001
DMULT2	0.0187	0.0001	0.0289	0.0001	-669.772	0.0001
DMULT3	0.0114	0.0001	0.0214	0.0001	-463.372	0.0001
DMULT4	0.0212	0.0196	0.0385	0.0001	-655.812	0.0007
NSTAY1	0.0022	0.0143	0.0168	0.0001	247.631	0.0001
NSTAY2	0.0021	0.1980	0.0151	0.0001	15.600	0.6572
NSTAY3	-0.0070	0.0068	0.0072	0.0054	-26.264	0.6348
POUT05	0.0422	0.0001	0.0511	0.0001	-1010.712	0.0001

Table 3.6

PATIENT CHARACTERISTIC MODELS, MEDICAL DRGS

	5% Longest Stays		5% Largest Losses		Accounting Profit	
Variable	Parameter Estimate	Significance Level	Parameter Estimate	Significance Level	Parameter Estimate	Significance Level
INTERCEP	-0.0410	0.0001	0.1260	0.0001	-1926.558	0.0001
UNDER 65	0.0186	0.1402	-0.1687	0.0001	4355.600	0.0001
LAGE	0.0162	0.0001	-0.0256	0.0001	603.431	0.0001
UNDER 65*LAGE	-0.0034	0.2617	0.0396	0.0001	-1040.565	0.0001
MALE	-0.0065	0.0001	0.0050	0.0001	22.829	0.0001
BLACK	0.0152	0.0001	0.0082	0.0001	-76.662	0.0001
RENDUM	-0.0161	0.0001	0.0137	0.0001	198.046	0.0001
OBSTRC	-0.0045	0.0001	0.0080	0.0001	-120.963	0.0001
ASTHMA	-0.0019	0.3180	-0.0001	0.9608	-139.047	0.0001
LFIB	-0.0057	0.0249	-0.0055	0.0157	162.691	0.0001
DEPRES	0.0184	0.0001	-0.0104	0.0001	-47.803	0.0199
LNGCAN	0.0093	0.0001	0.0083	0.0001	-38.033	0.0417
ATHERO	0.0082	0.0001	-0.0057	0.0001	-34.351	0.0020
PULEDM	0.0211	0.0001	0.0474	0.0001	-750.751	0.0001
LEUK	0.0181	0.0001	0.0297	0.0001	-249.721	0.0001
MALNUT	0.0384	0.0001	0.0416	0.0001	-476.753	0.0001
OBESE	-0.0148	0.0001	-0.0202	0.0001	195.455	0.0001
ANEMIA	-0.0024	0.0031	-0.0084	0.0001	141.521	0.0001
HHANEMIA	-0.0160	0.0127	-0.0086	0.1365	90.818	0.2719
ARTHR	-0.0161	0.0001	-0.0320	0.0001	356.036	0.0001
DIAB2	-0.0154	0.0001	-0.0236	0.0001	353.585	0.0001
DIAB3	-0.0122	0.0001	-0.0121	0.0001	296.232	0.0001
DIAB4	-0.0118	0.0001	-0.0146	0.0001	245.486	0.0001
DIAB5	-0.0028	0.0564	-0.0028	0.0365	56.309	0.0032
HYPER	-0.0148	0.0001	-0.0181	0.0001	276.895	0.0001
DEMENT	0.0042	0.0144	-0.0086	0.0001	136.734	0.0001
PSYCH	0.0087	0.0001	-0.0013	0.5175	19.930	0.4807
LIVER	0.0062	0.0052	0.0170	0.0001	-139.156	0.0001
DMULT1	0.0161	0.0001	0.0209	0.0001	-328.038	0.0001
DMULT2	0.0136	0.0001	0.0167	0.0001	-264.871	0.0001
DMULT3	0.0109	0.0001	0.0112	0.0001	-167.843	0.0001
DMULT4	0.0160	0.0001	0.0125	0.0001	-230.618	0.0001
NSTAY1	0.0067	0.0001	0.0045	0.0001	-20.356	0.0006
NSTAY2	-0.0012	0.0887	-0.0001	0.8610	17.799	0.0528
NSTAY3	-0.0078	0.0001	-0.0052	0.0001	95.246	0.0001
POUT05	0.0452	0.0001	0.0424	0.0001	-645.701	0.0001

who have had a prior expensive or long stay or who have multiple diseased body systems. The multiple diseased body system variables are additive. For example, for surgical cases (Table 3.5), the involvement of more than two body systems adds 0.0187 above the 0.0355 contribution of having more than one diseased body system. So the total difference in the likelihood of a long stay between a patient with more than two diseased body systems and a patient with only one diseased body system is 0.0522. For medical DRGs, malnutrition is the only secondary condition that substantially raises the likelihood of a long stay; multiple diseased body systems and a prior long or expensive stay also result in large increases in the probability of a long stay.

The likelihood of an extremely large loss increases most for patients who have a secondary condition of pulmonary edema and malnutrition, have multiple system involvement, or have had a prior expensive or long stay. For surgical cases, the likelihood of an extremely large loss also increases for patients with secondary conditions of lung cancer and chronic liver disease. Patients in surgical DRGs with these characteristics have a greater likelihood of being extremely expensive than do those patients in medical DRGs. That is, although malnutrition increases the likelihood of being an expensive case for both surgical and medical patients, a much greater likelihood exists that a surgical patient with malnutrition will become an extremely expensive case than a medical patient.

In the third equation, negative coefficients mean that a typical patient with that characteristic is less profitable than average and is likely to cause losses for many hospitals. Cases with pulmonary edema or malnutrition, or a previous long or expensive stay, cause substantial losses for hospitals. Involvement of multiple body systems also frequently translates to a large loss. Although two or more diseased body systems for a surgical case lead to additional losses of $798, three or more combine to an additional loss of $670. For medical cases, having leukemia is also associated with large losses. For surgical cases, being black is associated with large losses. Surgical cases are more profitable on average. However, surgical patients with secondary conditions of pulmonary edema or malnutrition lead to much larger losses on average than do medical patients.

COMPARISON OF EXTREMELY LONG STAYS AND EXTREMELY LARGE LOSSES

HCFA has two criteria for paying outliers. The first criterion identifies cases based on a length of stay that exceeds an absolute threshold, or based on a long length of stay relative to the average length of stay for the DRG. A second criterion based on a case's cost is used if the length of stay criterion is not met. A cost outlier identifies those cases having a relatively short stay but whose resource requirements are exceedingly high either in absolute terms or relative to the DRG. As discussed earlier, HCFA has modified the outlier policy to pay the maximum of the day or cost outlier payment.

Because payment policies differ according to day and cost criteria, we were interested in seeing if certain patient characteristics were more closely linked to longer stays or to larger losses. Comparisons between the models for extremely long and extremely costly stays can be made by referring back to Tables 3.5 and 3.6.

Most coefficients for the patient characteristic variables exhibit the same sign. Characteristics positively associated with the longest stays are also positively associated with the largest losses. Two secondary conditions—end-stage renal disease and depression—have noteworthy reversals in direction between long stays and large losses. For cases in medical DRGs, a secondary condition of end-stage renal disease is positively associated with the largest losses, but is negatively associated with the longest stays. A secondary condition of depression for medical cases, on the other hand, is positively associated with the longest stays but negatively associated with the largest losses.

EFFECTS OF AGE AND DISABILITY

Normally, we would expect to find increasing severity with advancing age. However, the Medicare population includes disabled persons in addition to the aged. Medicare patients under 65 years of age are all disabled and may be more expensive to care for on average than those 65 years and over (although the aged group also includes some disabled persons). Because of this change in populations at age 65, we included in our models an indicator for patients under 65 years and an interaction of log of age and the indicator, in addition to log of age. Instead of interpreting the three regression coefficients associated with these three age variables directly, it helps to calculate the combined contribution at given ages. The contribution of age for persons under 65 years is the sum of the coefficient for the indicator for those under 65 years, the coefficient for the interaction term multiplied by the log of age, and the coefficient for LAGE multipled by the log of age. For patients 65 years and older, the contribution of age is only the coefficient for LAGE multiplied by the log of age. Table 3.7 shows these calculations for the longest stays, the largest losses, and the profit models for cases in surgical DRGs.

To better demonstrate the effects of disability and increasing age on the likelihood of long or expensive stays, we use age 65 as a comparison group. Tables 3.8 and 3.9 show the calculated contribution for each age group less the contribution of being age 65. For instance,

Table 3.7

COMBINED EFFECT OF THE AGE COEFFICIENTS, SURGICAL DRGS

Age	Length of Stay	Extreme Losses	Profit
Disabled only			
40	.4809	.2326	-6725.68
50	.4827	.2281	-6714.01
60	.4841	.2244	-6704.48
Aged and/or disabled			
65	.4784	.2179	-6714.00
70	.4869	.2218	-6833.19
80	.5022	.2287	-7047.96
90	.5157	.2349	-7237.40

in Table 3.8, an individual 90 years of age has a calculated coefficient 0.0373 greater than does an individual 65 years of age. So within the aged population, increasing age is associated with an increasing likelihood of becoming a long stay.

For aged patients (65 and over) in surgical DRGs (Table 3.8), increasing age is associated with an increasing likelihood of becoming an extremely long stay and an extremely expensive stay, but is associated with decreasing profit. For disabled patients in surgical DRGs, increasing age is associated with negligible changes in the likelihood of becoming an extremely long stay, and a slightly decreasing likelihood of becoming an extremely expensive stay. Being disabled increases the likelihood of a long or expensive stay and is associated with decreased profit.

For both disabled and aged patients in medical DRGs, increasing age is associated with small increases in the likelihood of becoming an extremely long stay, but decreasing likelihood of becoming an extremely expensive stay. For long stays, no significant difference between the disabled and aged populations exists; the disabled population, however, is less apt to have a large loss.

COLLINEARITY OF MULTIPLE DISEASED BODY SYSTEMS AND CHRONIC DISEASES

In general, the inclusion or exclusion of sets of explanatory variables in regression models can affect the coefficients for other variables in the model. During exploratory analysis we noted that including or excluding the demographic or prior utilization variables had little impact on the coefficients in the models. However, the chronic disease and multiple system variables do affect each other's coefficients. We should expect this finding given the construction of the variables.

Because we can envision two separate policies—one aimed at multiple diseased body systems, the other at specific chronic diseases—exploring their interaction a bit further is worthwhile. Tables 3.10 and 3.11 show the model explaining large losses, with the chronic disease variables omitted for surgical and medical cases, respectively. The

Table 3.8

COMPARISON OF AGE CONTRIBUTION TO AGE 65, SURGICAL DRGS

Age	Length of Stay	Extreme Losses	Profit
Disabled only			
40	.0076	.0147	-11.69
50	.0093	.0102	-0.02
60	.0107	-.0065	9.52
Aged and/or disabled			
65	.0000	.0000	0.00
70	.0085	.0039	-119.19
80	.0238	.0108	-333.60
90	.0373	.0170	-523.40

Table 3.9

COMPARISON OF AGE CONTRIBUTION TO AGE 65, MEDICAL DRGS

Age	Length of Stay	Extreme Losses	Profit
Disabled only			
40	-.0018	-.0102	224.12
50	.0011	-.0070	126.57
60	.0034	-.0045	46.87
Aged and/or disabled			
65	.0000	.0000	0.00
70	.0012	-.0019	44.72
80	.0034	-.0053	125.30
90	.0053	-.0083	196.37

last model in each table repeats the coefficients of the full model explained earlier. As can be seen, the coefficients for multiple diseased body systems are larger with the full model. In addition, as compared to the model with the chronic disease variables only, the full model shows that the effects of many chronic disease variables diminish and sometimes even change sign.

The existence of a secondary chronic disease almost always causes an increase in the count of diseased body systems. However, because mental illness and neoplasms are not included in our count of multiple diseased body systems, five of our chronic variables remain unaffected. Lung cancer, leukemia, depression, psychosis, and dementia have coefficients that are unaffected by the addition of the multiple system variables.

The collinearity is strong enough to affect the coefficients substantially. The effects of many individual chronic conditions can be captured by the counts of the number of involved body systems. However, a few of the individual chronic diseases—pulmonary edema, malnutrition, and liver disease—increase the chance of a very costly case noticeably beyond the increase caused by the additional involved body system. This correlation holds especially true for surgical patients. Thus, if one did wish to vary payments based on patient characteristics that increase the risk of a very expensive case, then the minimal set of characteristics to consider would include the number of involved body systems, chronic diseases not counted as involved body systems, and these three diseases. However, determining that these conditions were not a complication of therapy and the result of poor quality of care would be important.

CONTROLLING FOR HOSPITAL

If certain patient groups tend to use hospitals that on average incur more losses, the patient characteristic may be picking up an effect that is really attributable to the hospital. To control for the hospital's effect, we ran additional regression models that included a dummy variable for each hospital. Although the coefficients for each

Table 3.10

PATIENT CHARACTERISTIC MODELS FOR 5 PERCENT LARGEST LOSSES, SURGICAL DRGS

	Excludes Chronic Diseases		Excludes Multiple Body Systems		Full Model	
Variable	Parameter Estimate	Significance Level	Parameter Estimate	Significance Level	Parameter Estimate	Significance Level
INTERCEP	-0.2645	0.0001	-0.2709	0.0001	-0.1950	0.0001
UNDER 65	0.4061	0.0001	0.4023	0.0001	0.3071	0.0001
LAGE	0.0673	0.0001	0.0730	0.0001	0.0522	0.0001
UNDER 65*LAGE	-0.0957	0.0001	-0.0944	0.0001	-0.0724	0.0001
MALE	0.0087	0.0001	0.0070	0.0001	0.0053	0.0001
BLACK	0.0216	0.0001	0.0252	0.0001	0.0227	0.0001
NSTAY1	0.0182	0.0001	0.0176	0.0001	0.0168	0.0001
NSTAY2	0.0169	0.0001	0.0168	0.0001	0.0151	0.0001
NSTAY3	0.0107	0.0001	0.0088	0.0006	0.0072	0.0054
POUT05	0.0541	0.0001	0.0550	0.0001	0.0511	0.0001
DMULT1	0.0334	0.0001			0.0472	0.0001
DMULT2	0.0181	0.0001			0.0289	0.0001
DMULT3	0.0149	0.0001			0.0214	0.0001
DMULT4	0.0339	0.0002			0.0385	0.0001
RENDUM			0.0141	0.0001	0.0034	0.2917
OBSTRC			0.0151	0.0001	-0.0187	0.0001
ASTHMA			-0.0036	0.4487	-0.0319	0.0001
LFIB			-0.0036	0.6028	-0.0300	0.0001
DEPRES			0.0119	0.0356	0.0083	0.1399
LNGCAN			0.0419	0.0001	0.0497	0.0001
ATHERO			0.0218	0.0001	0.0098	0.0001
PULEDM			0.2328	0.0001	0.1990	0.0001
LEUK			0.0217	0.0077	0.0206	0.0109
OBESE			-0.0120	0.0001	-0.0399	0.0001
MALNUT			0.1718	0.0001	0.1382	0.0001
ANEMIA			0.0191	0.0001	-0.0166	0.0001
HHANEMIA			-0.0056	0.7334	-0.0367	0.0259
ARTHR			-0.0161	0.0001	-0.0436	0.0001
DIAB2			-0.0104	0.0001	-0.0440	0.0001
DIAB3			0.0091	0.0001	-0.0251	0.0001
DIAB4			0.0181	0.0001	-0.0205	0.0001
DIAB5			0.0434	0.0001	0.0038	0.2407
HYPER			-0.0225	0.0001	-0.0437	0.0001
DEMENT			-0.0165	0.0006	-0.0172	0.0003
PSYCH			0.0074	0.2677	0.0083	0.2126
LIVER			0.0783	0.0001	0.0586	0.0001
R-SQUARE	0.0173		0.0162		0.0276	

Table 3.11

PATIENT CHARACTERISTIC MODELS FOR 5 PERCENT LARGEST LOSSES, MEDICAL DRGS

	Excludes Chronic Diseases		Excludes Multiple Body Systems		Full Model	
Variable	Parameter Estimate	Significance Level	Parameter Estimate	Significance Level	Parameter Estimate	Significance Level
INTERCEP	0.0943	0.0001	0.0611	0.0001	0.1260	0.0001
UNDER 65	-0.1218	0.0001	-0.1064	0.0001	-0.1687	0.0001
LAGE	-0.0191	0.0001	-0.0077	0.0001	-0.0256	0.0001
UNDER 65*LAGE	0.0285	0.0001	0.0251	0.0001	0.0396	0.0001
MALE	0.0080	0.0001	0.0043	0.0001	0.0050	0.0001
BLACK	0.0067	0.0001	0.0092	0.0001	0.0082	0.0001
NSTAY1	0.0061	0.0001	0.0050	0.0001	0.0045	0.0001
NSTAY2	0.0009	0.1640	0.0001	0.8582	-0.0001	0.8610
NSTAY3	-0.0042	0.0001	-0.0062	0.0001	-0.0052	0.0001
POUT05	0.0437	0.0001	0.0445	0.0001	0.0424	0.0001
DMULT1	0.0165	0.0001			0.0209	0.0001
DMULT2	0.0128	0.0001			0.0167	0.0001
DMULT3	0.0083	0.0001			0.0112	0.0001
DMULT4	0.0091	0.0015			0.0125	0.0001
RENDUM			0.0246	0.0001	0.0137	0.0001
OBSTRC			0.0184	0.0001	0.0080	0.0001
ASTHMA			0.0075	0.0001	-0.0001	0.9608
LFIB			0.0010	0.6622	-0.0055	0.0157
DEPRES			-0.0110	0.0001	-0.0104	0.0001
LNGCAN			0.0020	0.1203	0.0083	0.0001
ATHERO			-0.0003	0.6561	-0.0057	0.0001
PULEDM			0.0587	0.0001	0.0474	0.0001
LEUK			0.0281	0.0001	0.0297	0.0001
OBESE			-0.0097	0.0001	-0.0202	0.0001
MALNUT			0.0515	0.0001	0.0416	0.0001
ANEMIA			0.0078	0.0001	-0.0084	0.0001
HHANEMIA			0.0032	0.5768	-0.0086	0.1365
ARTHR			-0.0174	0.0001	-0.0320	0.0001
DIAB2			-0.0098	0.0001	-0.0236	0.0001
DIAB3			0.0021	0.0403	-0.0121	0.0001
DIAB4			0.0014	0.3238	-0.0146	0.0001
DIAB5			0.0139	0.0001	-0.0028	0.0365
HYPER			-0.0134	0.0001	-0.0181	0.0001
DEMENT			-0.0083	0.0001	-0.0086	0.0001
PSYCH			-0.0024	0.2240	-0.0013	0.5175
LIVER			0.0279	0.0001	0.0170	0.0001
R-SQUARE	0.0089		0.0087		0.0144	

hospital are not particularly of interest, including the hospital dummy variables will pick up the association of any hospital with the largest losses.

We reran the regressions for the top 5 percent of losses and length of stay controlling for the effect of individual hospitals. We ran these regressions separately for cases in surgical and medical DRGs. Tables 3.12 through 3.15 present the results of these analyses. The first model is the same model presented in earlier discussions, in which there is no control for hospital. The second model presents the coefficients for the patient characteristics in which dummies for each hospital have been included.

Almost all the coefficients change very little when the hospital's effect is added to the regression equation. This lack of significant change implies that these clinical characteristics predict cases that will have large losses no matter what hospital is involved—that is, these clinical characteristics do not simply occur in cases at expensive hospitals.

Race is one of the few exceptions. After controlling for hospital, race has less explanatory power for determining large losses. Although a significant, positive effect still remains, it is much smaller after controlling for the hospital. This finding implies that blacks are more likely to be admitted to a hospital that is more expensive, such as a disproportionate-share hospital.

Controlling for the hospital did not reduce the explanatory power for the multisystem disease variables. It reduced the effect of the prior utilization variables, although a previous expensive or long length of stay remains one of the large effects in the model.

Table 3.12

PATIENT CHARACTERISTIC MODELS FOR 5 PERCENT LARGEST LOSSES WITH AND WITHOUT CONTROL VARIABLES FOR EACH HOSPITAL, SURGICAL DRGS

	Not Controlling for Hospital		Controlling for Hospital	
Variable	Parameter Estimate	Significance Level	Parameter Estimate	Significance Level
INTERCEP	-0.1950	0.0001	(a)	(a)
UNDER 65	0.3071	0.0001	0.2946	0.0001
LAGE	0.0522	0.0001	0.0566	0.0001
UNDER 65*LAGE	-0.0724	0.0001	-0.0697	0.0001
MALE	0.0053	0.0001	0.0048	0.0001
BLACK	0.0227	0.0001	0.0141	0.0001
RENDUM	0.0034	0.2917	-0.0034	0.2712
OBSTRC	-0.0187	0.0001	-0.0168	0.0001
ASTHMA	-0.0319	0.0001	-0.0299	0.0001
LFIB	-0.0300	0.0001	-0.0273	0.0001
DEPRES	0.0083	0.1399	0.0084	0.1127
LNGCAN	0.0497	0.0001	0.0467	0.0001
ATHERO	0.0098	0.0001	0.0084	0.0001
PULEDM	0.1990	0.0001	0.1857	0.0001
LEUK	0.0206	0.0109	0.0187	0.0147
MALNUT	0.1382	0.0001	0.1299	0.0001
OBESE	-0.0399	0.0001	-0.0363	0.0001
ANEMIA	-0.0166	0.0001	-0.0159	0.0001
HHANEMIA	-0.0367	0.0259	-0.0409	0.0092
ARTHR	-0.0436	0.0001	-0.0397	0.0001
DIAB2	-0.0440	0.0001	-0.0413	0.0001
DIAB3	-0.0251	0.0001	-0.0241	0.0001
DIAB4	-0.0205	0.0001	-0.0203	0.0001
DIAB5	0.0038	0.2407	0.0037	0.2262
HYPER	-0.0437	0.0001	-0.0424	0.0001
DEMENT	-0.0172	0.0003	-0.0153	0.0007
PSYCH	0.0083	0.2126	0.0051	0.4234
LIVER	0.0586	0.0001	0.0528	0.0001
DMULT1	0.0472	0.0001	0.0450	0.0001
DMULT2	0.0289	0.0001	0.0280	0.0001
DMULT3	0.0214	0.0001	0.0194	0.0001
DMULT4	0.0385	0.0001	0.0366	0.0001
NSTAY1	0.0168	0.0001	0.0146	0.0001
NSTAY2	0.0151	0.0001	0.0137	0.0001
NSTAY3	0.0072	0.0054	0.0076	0.0018
POUT05	0.0511	0.0001	0.0390	0.0001

(a) The computational technique that controls for the hospital omits the coefficients for the hospital dummies and the intercept term.

Table 3.13

PATIENT CHARACTERISTIC MODELS FOR 5 PERCENT LARGEST LOSSES WITH AND WITHOUT CONTROL VARIABLES FOR EACH HOSPITAL, MEDICAL DRGS

	Not Controlling for Hospital		Controlling for Hospital	
Variable	Parameter Estimate	Significance Level	Parameter Estimate	Significance Level
INTERCEP	0.1260	0.0001	(a)	(a)
UNDER 65	-0.1687	0.0001	-0.1542	0.0001
LAGE	-0.0256	0.0001	-0.0183	0.0001
UNDER 65*LAGE	0.0396	0.0001	0.0364	0.0001
MALE	0.0050	0.0001	0.0048	0.0001
BLACK	0.0082	0.0001	0.0022	0.0006
RENDUM	0.0137	0.0001	0.0025	0.1049
OBSTRC	0.0080	0.0001	0.0085	0.0001
ASTHMA	-0.0001	0.9608	0.0003	0.8516
LFIB	-0.0055	0.0157	-0.0043	0.0435
DEPRES	-0.0104	0.0001	-0.0091	0.0001
LNGCAN	0.0083	0.0001	0.0053	0.0001
ATHERO	-0.0057	0.0001	-0.0041	0.0001
PULEDM	0.0474	0.0001	0.0465	0.0001
LEUK	0.0297	0.0001	0.0269	0.0001
MALNUT	0.0416	0.0001	0.0390	0.0001
OBESE	-0.0202	0.0001	-0.0176	0.0001
ANEMIA	-0.0084	0.0001	-0.0096	0.0001
HHANEMIA	-0.0086	0.1365	-0.0124	0.0208
ARTHR	-0.0320	0.0001	-0.0282	0.0001
DIAB2	-0.0236	0.0001	-0.0221	0.0001
DIAB3	-0.0121	0.0001	-0.0115	0.0001
DIAB4	-0.0146	0.0001	-0.0122	0.0001
DIAB5	-0.0028	0.0365	-0.0014	0.2687
HYPER	-0.0181	0.0001	-0.0184	0.0001
DEMENT	-0.0086	0.0001	-0.0086	0.0001
PSYCH	-0.0013	0.5175	-0.0018	0.3400
LIVER	0.0170	0.0001	0.0133	0.0001
DMULT1	0.0209	0.0001	0.0206	0.0001
DMULT2	0.0167	0.0001	0.0159	0.0001
DMULT3	0.0112	0.0001	0.0105	0.0001
DMULT4	0.0125	0.0001	0.0117	0.0001
NSTAY1	0.0045	0.0001	0.0044	0.0001
NSTAY2	-0.0001	0.8610	0.0006	0.2971
NSTAY3	-0.0052	0.0001	-0.0039	0.0001
POUT05	0.0424	0.0001	0.0311	0.0001

(a) The computational technique that controls for the hosptial omits the coefficients for the hospital dummies and the intercept term.

Table 3.14

PATIENT CHARACTERISTIC MODELS FOR 5 PERCENT LONGEST STAYS WITHIN DRG WITH OR WITHOUT CONTROL VARIABLES FOR EACH HOSPITAL, SURGICAL DRGS

	Not Controlling for Hospital		Controlling for Hospital	
Variable	Parameter Estimate	Significance Level	Parameter Estimate	Significance Level
INTERCEP	-0.4523	0.0001	(a)	(a)
UNDER 65	0.4572	0.0001	0.4812	0.0001
LAGE	0.1146	0.0001	0.1258	0.0001
UNDER 65*LAGE	-0.1068	0.0001	-0.1127	0.0001
MALE	0.0035	0.0001	0.0042	0.0001
BLACK	0.0393	0.0001	0.0223	0.0001
RENDUM	-0.0281	0.0001	-0.0282	0.0001
OBSTRC	-0.0318	0.0001	-0.0302	0.0001
ASTHMA	-0.0116	0.0167	-0.0087	0.0615
LFIB	-0.0268	0.0001	-0.0239	0.0005
DEPRES	0.0528	0.0001	0.0532	0.0001
LNGCAN	0.0499	0.0001	0.0491	0.0001
ATHERO	0.0094	0.0001	0.0085	0.0001
PULEDM	0.0641	0.0001	0.0597	0.0001
LEUK	0.0060	0.4602	0.0048	0.5472
MALNUT	0.0458	0.0001	0.0458	0.0001
OBESE	-0.0262	0.0001	-0.0232	0.0001
ANEMIA	-0.0165	0.0001	-0.0129	0.0001
HHANEMIA	-0.0532	0.0014	-0.0476	0.0034
ARTHR	-0.0124	0.0001	-0.0108	0.0001
DIAB2	-0.0195	0.0001	-0.0213	0.0001
DIAB3	-0.0222	0.0001	-0.0231	0.0001
DIAB4	-0.0075	0.0409	-0.0083	0.0213
DIAB5	-0.0073	0.0248	0.0055	0.0797
HYPER	-0.0296	0.0001	-0.0301	0.0001
DEMENT	0.0021	0.6570	-0.0049	0.2898
PSYCH	0.0307	0.0001	0.0283	0.0001
LIVER	0.0119	0.0084	0.0105	0.0162
DMULT1	0.0355	0.0001	0.0364	0.0001
DMULT2	0.0187	0.0001	0.0183	0.0001
DMULT3	0.0114	0.0001	0.0123	0.0001
DMULT4	0.0212	0.0196	0.0240	0.0069
NSTAY1	0.0022	0.0143	0.0015	0.0986
NSTAY2	0.0021	0.1980	0.0026	0.0986
NSTAY3	-0.0070	0.0068	-0.0072	0.0043
POUT05	0.0422	0.0001	0.0326	0.0001

(a) The computational technique that controls for the hosptial omits the coefficients for the hospital dummies and the intercept term.

Table 3.15

PATIENT CHARACTERISTIC MODELS FOR 5 PERCENT LONGEST STAYS WITHIN DRG WITH OR WITHOUT CONTROL VARIABLES FOR EACH HOSPITAL, MEDICAL DRGS

	Not Controlling for Hospital		Controlling for Hospital	
Variable	Parameter Estimate	Significance Level	Parameter Estimate	Significance Level
INTERCEP	-0.0410	0.0001	(a)	(a)
UNDER 65	0.0186	0.1402	0.0662	0.0001
LAGE	0.0162	0.0001	0.0303	0.0001
UNDER 65*LAGE	-0.0034	0.2617	-0.0148	0.0001
MALE	-0.0065	0.0001	-0.0061	0.0001
BLACK	0.0152	0.0001	0.0043	0.0001
RENDUM	-0.0161	0.0001	-0.0222	0.0001
OBSTRC	-0.0045	0.0001	-0.0036	0.0001
ASTHMA	-0.0019	0.3180	0.0007	0.6899
LFIB	-0.0057	0.0249	-0.0069	0.0048
DEPRES	0.0184	0.0001	0.0182	0.0001
LNGCAN	0.0093	0.0001	0.0062	0.0001
ATHERO	0.0082	0.0001	0.0085	0.0001
PULEDM	0.0211	0.0001	0.0204	0.0001
LEUK	0.0181	0.0001	0.0191	0.0001
MALNUT	0.0384	0.0001	0.0385	0.0001
OBESE	-0.0148	0.0001	-0.0128	0.0001
ANEMIA	-0.0024	0.0031	-0.0036	0.0001
HHANEMIA	-0.0160	0.0127	-0.0205	0.0008
ARTHR	-0.0161	0.0001	-0.0155	0.0001
DIAB2	-0.0154	0.0001	-0.0170	0.0001
DIAB3	-0.0122	0.0001	-0.0130	0.0001
DIAB4	-0.0118	0.0001	-0.0088	0.0001
DIAB5	-0.0028	0.0564	-0.0013	0.3735
HYPER	-0.0148	0.0001	-0.0157	0.0001
DEMENT	0.0042	0.0144	0.0045	0.0064
PSYCH	0.0087	0.0001	0.0093	0.0001
LIVER	0.0062	0.0052	0.0050	0.0190
DMULT1	0.0161	0.0001	0.0172	0.0001
DMULT2	0.0136	0.0001	0.0140	0.0001
DMULT3	0.0109	0.0001	0.0108	0.0001
DMULT4	0.0160	0.0001	0.0156	0.0001
NSTAY1	0.0067	0.0001	0.0066	0.0001
NSTAY2	-0.0012	0.0887	0.0001	0.8540
NSTAY3	-0.0078	0.0001	-0.0066	0.0001
POUT05	0.0452	0.0001	0.0359	0.0001

(a) The computational technique that controls for the hospital omits the coefficients for the hospital dummies and the intercept term.

IV. DISCUSSION

With the advent of PPS came concern that the more severely ill patient could potentially be denied access to care or be provided with less care than might be necessary. Berenson and Pawlson (1984) argue that if specific types of patients, such as the frail elderly, are revenue losers for hospitals, hospitals may systematically try to limit admissions of these patient groups—or to limit the provision of care to these patients.

Limiting access to care might occur in more subtle ways than blatantly discriminatory admission policies or hospital administrators' cajoling physicians to admit costly patients elsewhere. Hospitals might choose to reduce or discontinue services predominantly used by sicker and costlier patients. Stern and Epstein (1985) suggest that hospitals might select staff based on the attractiveness of the physicians' clientele. Thus, problems of access can be viewed not only as turning away certain patients at the door, but also as hospitals' designing a mix of services favoring more profitable patient groups.

Newhouse (1989) investigated problems of access for unprofitable patients by looking at transfers and shifts in hospital caseload. He did not find that voluntary and proprietary hospitals had transferred unprofitable patients to hospitals of last resort. Since peer review organizations monitor hospital transfers, we should not be surprised by this finding. That hospitals would fail to admit unprofitable patient groups entirely (rather than admit them and then transfer them) is more likely, since many attributes associated with an unprofitable patient are observable before admission. For instance, if the admitting physician is the patient's primary care physician, that physician would be well aware of previous hospitalizations and chronic disease conditions.

Newhouse (1989) did find more subtle forms of selection, demonstrated by a systematic tendency for unprofitable DRGs to be a higher share of the caseload in hospitals of last resort. This finding is consistent with limiting access—for example, by choosing not to provide the types of care associated with unprofitable patients.

In the earlier years of PPS, most hospitals profited under DRG reimbursement. As adjustments have been made over time, hospitals are facing more and more fiscal constraints; as a result, they have more incentive to monitor physicians' admissions and costs, as well as to limit resources available to costlier patient groups. Since Newhouse (1989) used 1984 Medicare data and found a shift to hospitals of last resort, we would expect to find a larger shift using 1988 data. In addition, we might find a larger share of the costlier patient groups being admitted to hospitals of last resort. The patient groups that had lower than average profits in 1985 probably incurred substantial losses in 1988.

After deciding that certain types of patients might be at risk for limited access to care, we then want to determine which particular groups are at risk and what might be done to protect them. Jencks and Kay (1987) chose eight patient groups and considered the value of further refining the DRG payment system based on these patient characteristics. The patient groups they investigated were nursing home residents, prior hospitalizations (within one year), the poor, the aged and very aged, disabled beneficiaries under age 65, end-stage renal disease beneficiaries under age 65, and disabled beneficiaries over age 65. They were specifically interested in whether part of the variation within DRGs could be attributed to any of these patient groups. They found that patients who had been in a nursing home before their hospitalization had charges 6.2 percent higher than beneficiaries who had not been nursing home residents. They also found that beneficiaries hospitalized previously within the calendar year had charges 3.2 percent higher. Older disabled beneficiaries had charges 6.7 percent higher than other beneficiaries. They concluded that substantial refinements could not be made based on these patient characteristics. They further concluded that hospital administrators could not efficiently discriminate against these specific patient groups and would prefer admitting a costlier patient to leaving a hospital bed empty.

Admittedly, discriminating against a large class of patients is hard. However, potential exists for discrimination against a smaller patient group associated with large losses. We found that 38 percent of the patients in this study had been hospitalized six months before their current hospital stay. This patient group is too large for hospital administrators to avoid; this group also does not represent substantial losses. However, if we consider previous hospitalizations that resulted in a large loss or a long stay relative to the DRG, we are now considering a population that is only 4 percent of the beneficiary hospital population and that represents both larger losses on average and an increased likelihood of substantial loss. Some 4.6 percent of the patients have more than three diseased body systems, and approximately 0.3 percent of the patients have pulmonary edema. These groups are also small and represent substantial losses for the hospital.

Note that since physicians admit patients to hospitals, they may discriminate on a case-by-case basis rather than against a particular patient group. A physician would be familiar with a patient's previous hospitalizations and their associated costs and would have a general impression of the current stay's cost. Thus, hospitals have incentives to persuade physicians to admit or refer costlier patients to hospitals of last resort.

Guterman (1986) identified several patient types that tend to become outliers. He looked at beneficiaries by age, end-stage renal disease, disability status, and discharge destination. He found that end-stage renal disease patients were more likely to become length of stay outliers and cost outliers, and that they also had higher average costs and charges.

Our study extends the work by Guterman (1986). Our methods differ in that rather than identifying cases that are reimbursed as outliers under current policy, we identify the 5 percent absolute largest losses and the 5 percent tail of long lengths of stay for DRGs. We found that specific patient characteristics do predict extremely long stays and large losses. Having a prior long or expensive stay, multiple diseased body systems, chronic liver disease, pulmonary edema, or leukemia is strongly associated with large losses. By identifying costlier patient groups, one underscores the necessity of outlier payments. Costlier care cannot be attributed solely to hospital inefficiency—it must in some part be attributable to patient need. If hospitals are not adequately reimbursed for caring for costlier patients, the hospitals that care for a greater number of these patients will either have to reduce resources across the board or for these costlier patient groups.

Before the beginning of fiscal year 1989, outlier policy paid for a case as a day outlier even if it met the criterion of a cost outlier as well. For some very costly patients, this policy did not reimburse appropriately. The outlier payment covered only the marginal per diem rate for the additional days beyond the day outlier threshold; then the reimbursement was not adequate to cover the more intense resource use also required by the patient during the stay. Hospitals treating patient groups that tended to have both longer stays and higher costs might not have been adequately compensated. Keeler et al. (1988) found that a substantial number of cases that exceeded the cost outlier threshold still had large losses because they were reimbursed by a small day outlier payment.

In fiscal year 1989, the HCFA began reimbursing the maximum of the day outlier payment and the cost outlier payment. Such a policy reduces the losses associated with the most unprofitable patients. However, Keeler et al. (1988) showed it does not reduce the losses associated with some groups of patients defined by clinical and demographic characteristics. The patient groups they considered were end-stage renal patients, diabetics with complications, blacks, and patients with previous long or expensive stays. Except for the end-stage renal disease patients, the average outlier payment for these groups was smaller because of the cost outlier threshold's outward shift. Similarly, except for the end-stage renal disease patients, the average profit per case was less. They then looked at the 5 percent and 10 percent most costly cases within each patient group. The average losses for the most costly cases increased for blacks and diabetic patients with complications. The average losses decreased for end-stage renal disease patients and patients with previous long or expensive stays. To accurately predict the effects by patient group of simultaneously shifting the cost threshold outward and paying the maximum of the day and cost outlier payment, the exact policy under consideration must be simulated.

Our study investigated both the effects of chronic conditions and the effects of having multiple diseased body systems. We found that the count of multiple body systems is easily defined and predicts substantially lower-than-average profits as well as an increased chance of a large loss. It also captures the effects of many individual chronic conditions. Three chronic diseases—pulmonary edema, malnutrition, and liver disease—increase the chance of a very costly stay beyond the increase caused by an additional involved body system. If one were to devise a payment system based on patient characteristics that increase the risk of a very costly case, one might well start with the count of diseased body systems. Further improvement would be gained by adding a small list of chronic conditions.

This research suggests that pulmonary congestion should be added to the list of secondary conditions and complications. Since it is currently not included in the list of conditions that result in an upcoding to a higher DRG payment, this change might reduce some of the large losses associated with this condition.

Appendix

COMPARISON OF 1988 AND 1986 DIAGNOSIS RELATED GROUPINGS

During the later stages of the analysis, the 1988 DRG Grouper became available. The 1988 version promised to be more specific and detailed. Earlier versions of the DRG grouping algorithm provided a list of diagnoses that would upcode to a DRG with complications and comorbidities. Any of the diagnoses on the list would result in a higher classification for each DRG "pair." With the 1988 version, some secondary diagnoses no longer applied to some of the principal diagnoses. For example, a secondary diagnosis of adult-onset diabetes with renal manifestations (25040) was no longer considered a complication or comorbidity for a primary diagnosis of hypoglycemic coma (2513x). In addition, being age 70 and over no longer affected the DRG classification.

With the changes from the 1986 version to the 1988 version of the DRGs, payment for DRGs with complications increased and those without decreased. That is, a larger payment differential between the DRG "pairs" accompanied the change that allowed for fewer complications and comorbidity. Thus, fewer cases now go to the higher of the DRG "pair," and these fewer cases are paid more.

Table A.1 shows an earlier generation of our patient characteristics model using the 1986 and 1988 versions of the DRG Grouper. As can be seen by Table A.1, the effect of all but four of the disease conditions decreased in absolute value from the 1986 version of DRGs to the 1988 version. Of these four, only juvenile-onset diabetes without complications and end-stage renal disease show changes of more than $20. And in these two cases, marginal losses of less than $100 were translated into marginal profits greater than $100.

Note that the end-stage renal disease condition is based on the Medicare disability status, not on the existence of a secondary diagnosis. Since the DRG Groupers do not take the Medicare status into account, the effect of the change for this condition is not readily predictable.

Table A.2 compares the 1986 and 1988 versions of Grouper for the 5 percent largest losses. All the coefficients for the disease conditions decreased in absolute value from the 1986 version to the 1988 version.

Table A.1

PROFIT

	DRG88		DRG86	
Variable	Parameter Estimate	Significance Level	Parameter Estimate	Significance Level
INTERCEP	2109.989	0.0001	1504.163	0.0001
UNDER 65	-64.294	0.7293	683.742	0.0003
LAGE	-365.102	0.0001	-195.372	0.0001
UNDER 65*LAGE	1.320	0.9766	-193.481	0.0001
MALE	-1.177	0.8342	13.213	0.0205
BLACK	-201.836	0.0001	-236.342	0.0001
RENDUM	156.665	0.0001	-39.299	0.1291
EMPHY	-120.989	0.0001	-351.307	0.0001
OBSTRC	-255.967	0.0001	-417.761	0.0001
ASTHMA	-257.223	0.0001	-286.265	0.0001
LFIB	63.982	0.1080	-221.681	0.0001
CEREB	-245.287	0.0001	-296.212	0.0001
DEPRES	-123.384	0.0001	-132.210	0.0001
LNGCAN	5.692	0.8057	-229.470	0.0001
ARTRY	159.143	0.0001	239.314	0.0001
PULEDM	-1192.363	0.0001	-1173.806	0.0001
LEUK	-230.176	0.0001	-538.795	0.0001
MALNUT	-383.649	0.0001	-558.543	0.0001
ANEMIA	-56.178	0.0001	-254.504	0.0001
ARTHR	127.884	0.0001	132.985	0.0001
DIAB2	165.591	0.0001	206.773	0.0001
DIAB3	128.822	0.0001	-56.367	0.0009
DIAB4	30.863	0.2023	-139.604	0.0001
DIAB5	-260.841	0.0001	-418.907	0.0001
HYPER	298.225	0.0001	370.542	0.0001
DEMENT	141.875	0.0001	122.556	0.0001
LIVER	-351.080	0.0001	-643.065	0.0001

These substantial changes in DRG coding therefore warranted use of the 1988 version of Grouper for this Note's final analysis.

Table A.2

FIVE PERCENT LARGEST LOSSES

	DRG88		DRG86	
Variable	Parameter Estimate	Significance Level	Parameter Estimate	Significcance Level
INTERCEP	0.0093	0.2210	-0.0188	0.0286
UNDER 65	-0.0057	0.5917	0.0228	0.0567
LAGE	0.0064	0.0003	0.0153	0.0001
UNDER 65*LAGE	0.0012	0.6343	-0.0050	0.0850
MALE	0.0063	0.0001	0.0070	0.0001
BLACK	0.0132	0.0001	0.0167	0.0001
RENDUM	0.0318	0.0001	0.0426	0.0001
EMPHY	0.0109	0.0001	0.0211	0.0001
OBSTRC	0.0168	0.0001	0.0264	0.0001
ASTHMA	0.0036	0.0315	0.0098	0.0001
LFIB	-0.0017	0.4716	0.0113	0.0005
CEREB	0.0060	0.0001	0.0093	0.0001
DEPRES	-0.0113	0.0001	-0.0124	0.0001
LNGCAN	0.0057	0.0001	0.0151	0.0001
ARTRY	-0.0003	0.7855	-0.0033	0.0141
PULEDM	0.0737	0.0001	0.0846	0.0001
LEUK	0.0259	0.0001	0.0368	0.0001
MALNUT	0.0245	0.0001	0.0357	0.0001
ANEMIA	0.0106	0.0001	0.0197	0.0001
ARTHR	-0.0199	0.0001	-0.0238	0.0001
DIAB2	-0.0115	0.0001	-0.0145	0.0001
DIAB3	0.0037	0.0001	0.0069	0.0001
DIAB4	0.0040	0.0043	0.0105	0.0001
DIAB5	0.0215	0.0001	0.0319	0.0001
HYPER	-0.0178	0.0001	-0.0243	0.0001
DEMENT	-0.0084	0.0001	-0.0080	0.0001
LIVER	0.0420	0.0001	0.0539	0.0001

REFERENCES

Berenson, Robert A., and L. Gregory Pawlson, "The Medicare Prospective Payment System and the Care of the Frail Elderly," *Journal of the American Geriatric Society,* Vol. 32, No. 11, November 1984, pp. 843-848.

Congressional Research Service, "Medicare's Prospective Payment System: An Analysis of the Financial Risk of Outlier Cases," Report 87-877 EPW, Library of Congress, Washington, D.C., 1987.

Federal Register, Vol. 52, No. 169, Tuesday, Sept. 1, 1987.

Federal Register, Vol. 53, No. 103, Friday, May 27, 1988.

Federal Register, Vol. 53, No. 190, Friday, Sept. 30, 1988.

Guterman, Stuart, "A Descriptive Analysis of PPS Outlier Payment Policy," Mimeo, 1986.

Haggstrom, Gus W., "Logistic Regression and Discriminant Analysis by Ordinary Least Squares," *Journal of Business and Economic Statistics*, Vol. 1, No. 3, July 1983, pp. 229-238.

Health Systems International, *DRGs Definitions Manual*, second and fourth revisions.

Jencks, Stephen F., and Terrence Kay, "Do Frail, Disabled, Poor, and Very Old Medicare Beneficiaries Have Higher Hospital Charges?" *Journal of the American Medical Association*, Vol. 257, 1987, pp. 198-202.

Keeler, Emmett B., Grace M. Carter, and Sally Trude, "Insurance Aspects of DRG Outlier Payments," The RAND Corporation, N-2762-HHS, 1988.

Mendenhall, S., "DRGs Must Be Changed to Take Patient's Illness Severity into Account,"*Modern Healthcare*, Vol. 14, No. 15, 1984, pp. 86-88.

Newhouse, Joseph P., "Do Unprofitable Patients Face Access Problems?" *HCFA Review*, 1989.

Stern, Robert S., and Arnold M. Epstein, "Institutional Responses to Prospective Payment Based on Diagnosis-Related Groups," *New England Journal of Medicine*, Vol. 312, No. 12, March 7, 1985, pp. 621-627.

U.S. Department of Health and Human Services, *The International Classification of Disease, 9th Revision, Clinical Modification*, Volumes 1 and 2, 2nd ed., Washington D.C., September 1980.

RAND/N-2998-HCFA

CMS LIBRARY
3 8095 00014377 2